Penetrometer Measurement of Apple and Kiwifruit Firmness: Operator and Instrument Differences

F.R. Harker[1], J.H. Maindonald, and P.J. Jackson
Horticulture and Food Research Institute of New Zealand, Mt. Albert Research Centre, Private Bag 92 169, Auckland, New Zealand

Additional index words. Instron, puncture, twist test, *Malus domestica, Actinidia deliciosa*

Abstract. Flesh firmness is a characteristic used to indicate fruit quality. Experimental design and data analysis are important when comparing devices that measure fruit firmness. We compared the Effegi penetrometer operated by hand, mounted in a drill press and then operated by hand, and mounted on a motorized drive and operated remotely; the hand-operated EPT pressure tester; the Instron with an Effegi probe; and a hand-operated prototype of the twist tester. Devices varied in operator differences and precision. Comparisons between devices were at the within-fruit level of variability and, therefore, more precise than comparisons where different device-operators used different fruit. We demonstrate statistical methods that are appropriate for making the comparisons of interest and discuss the possible cause of differences between operators and between devices. We also discuss how the mechanical properties of the devices may affect results and consider implications for their practical use. In this study, we found the precision of discrimination between soft and hard apples was best using the Instron in 1992, while the Instron and hand-held Effegi penetrometer were comparable in 1991. For kiwifruit, the hand-held Effegi penetrometer consistently gave the most precise measurements of softening in 1991, while the twist test was the most precise in 1992.

Fruit quality assessments usually involve measuring flesh firmness using a penetrometer. Generally, a probe, with either a flat or convex tip, is driven into the flesh, and the maximum force is recorded. Early hand-held penetrometers were developed by Magness and Taylor (1925), and from these devices, a range of devices have been developed by different companies (reviewed by Bourne, 1982). A number of studies have examined the mechanics of puncture testing (Voisey, 1977, and references therein) and recommendations on using penetrometers include those of Blanpied et al. (1978), Smith (1985), and Watkins and Harman (1981).

The wide use of firmness measurements has led to the periodic development of new devices for measuring firmness (e.g., Studman and Yuwana, 1992). Any adequate evaluation of their performance relative to existing equipment must consider operator differences, precision, and ability to detect texture change, as well as differences in measurement of absolute firmness. Effective experimental design is critical for comparing devices and comparing different ways of using the devices.

Previous studies have compared different types of penetrometers (e.g., Abbott et al., 1976; Bongers, 1992) and investigated how experimental design (replicate number and sample size) affects detection of firmness differences (Saltveit, 1978; Worthington and Yeatman, 1968). In our study, we used a balanced, incomplete blocks (paired comparison) experimental design together with modern statistical analysis approaches to compare different methods of measuring firmness. For relatively modest effort, our approach has provided information on differences associated with different operators and different methods and information on precision with which methods discriminate between soft and hard fruit. The usefulness of the design was evaluated by assessing the factor by which the number of measurements would need to be increased to achieve similar precision with a less controlled design.

Further design gains are possible. With large fruit, such as apples, as many as eight measurements per fruit may be possible without interference between results (Abbott et al., 1976). These multiple measurements would allow one pair of measurements for each of four device-operator combinations on each fruit. Biases that affect a particular device can, in principle, be corrected by a suitable calibration. Inherent lack of precision can be cured only by redesign of the device.

Materials and Methods

Plant material. In the 1991 and 1992 commercial apple harvests, fruit (*Malus domestica* L. cv. Granny Smith) harvested in Hawkes Bay were collected from the New Zealand Apple and Pear Marketing Board depot in Hastings and were transported to Auckland. Kiwifruit [*Actinidia deliciosa* (A. Chev) C.F. Liang et A.R. Ferguson cv. Hayward] were obtained from the HortResearch Orchard (Kumeu, Auckland) in both years. Kiwifruit and apples were placed in separate cool stores at 0 °C, and firmness experiments were undertaken within the first week and after storage. Apples were stored for 24 (1991) or 18 weeks (1992), and kiwifruit were stored for 17 (1991) or 15 (1992) weeks. All fruit were held at 20 °C for 1 d before firmness was measured.

Penetrometer devices. Measurements made in this study involve driving an Effegi probe with a convex tip into whole fruit. However, the mechanism used to drive the probe into the fruit differed. Probes with diameters of 11.1 and 7.9 mm were used to measure firmness of apples and kiwifruit, respectively. These probe sizes reflect standard methods for the respective fruit types.

The Effegi hand-held penetrometer (Facchini, Alfonsine, Italy) was the simplest of the devices (Fig. 1a). The fruit was held steady on a firm surface and the probe was pushed into the fruit to a depth of 8 mm, corresponding to a mark inscribed on the shaft of the probe (Watkins and Harman, 1981). In addition, the Effegi penetrometer was mounted either in a drill press (Black and Decker,
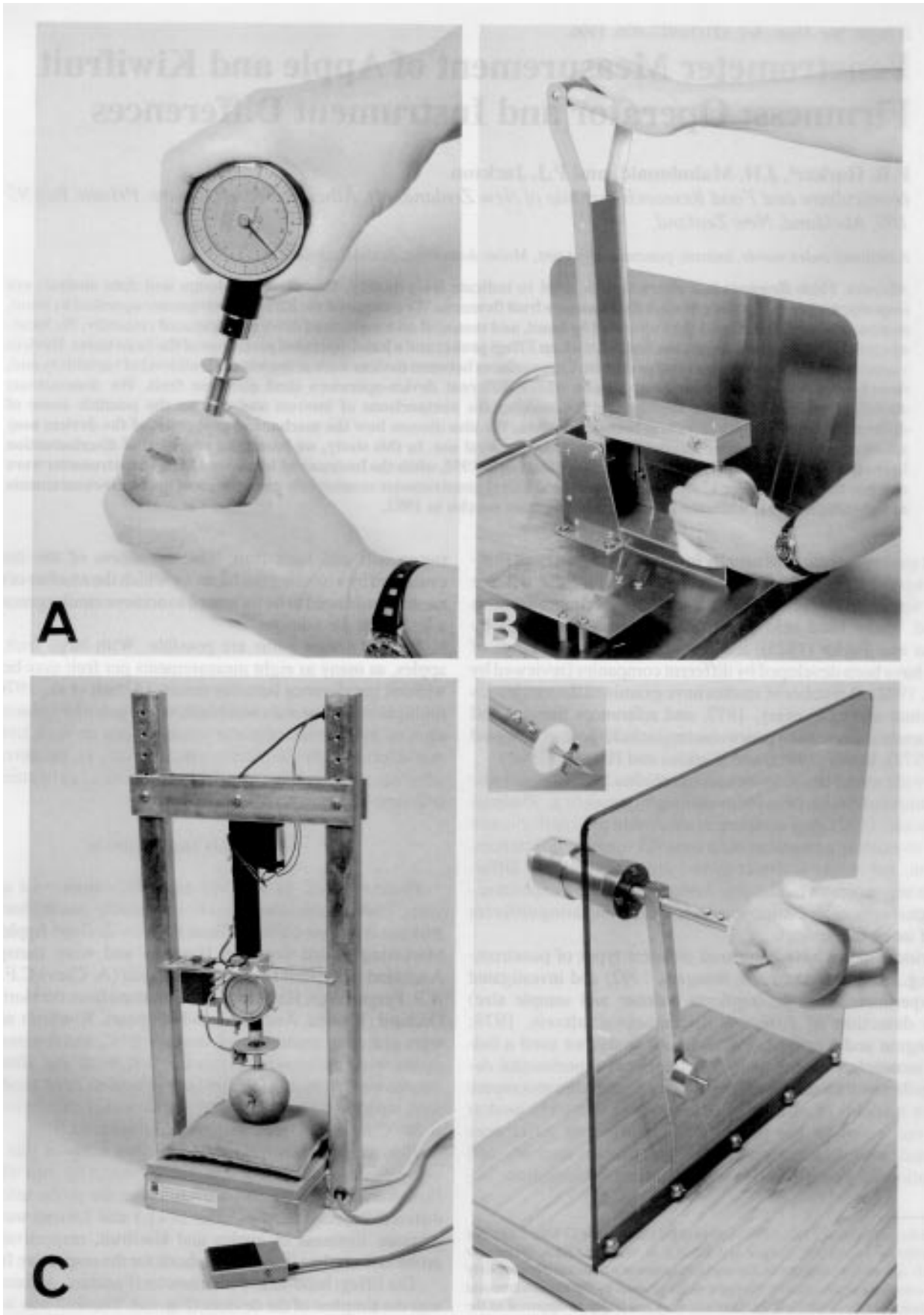
Fig. 1. Devices used to measure fruit firmness: (**A**) Effegi penetrometer, (**B**) EPT pressure tester, (**C**) Effegi penetrometer mounted in a motorized drive, and (**D**) the twist test (with insert showing blade). Instron and Effegi penetrometer mounted in drill press are not shown.

Hampstead, Md.) or a motorized press (Fig. 1c). The motorized press was developed by the Engineering Development Group of HortResearch as a standardized method of measuring kiwifruit firmness. The horizontal beam was driven downward at 1500 mm·min⁻¹. The end of the test was detected electronically when the probe had penetrated 8 mm into the fruit. The beam then returned automatically to its original position.

The EPT pressure tester (Lake City Technical Products, Kelowna, Canada; Fig. 1b) also is hand-operated. The penetrometer probe is located at the end of a pivoting arm, and the arm is pulled down until the probe enters the fruit to a depth of 8 mm. A spring returns the arm to its original position, and a load cell, located on a leaf spring within the arm, measures the force. The output is digital and can be imported into a computer or printed directly. The pressure tester interprets information on the shape of the penetration curve and provides messages when the measurement was taken "too fast" or "too slow", as well as the prompts "no slope" and "bad sample" when apples are overly soft or have inner pulpiness or bruises. The EPT pressure tester was used in the "MT off" mode.

Materials testing machines are widely used in assessing texture in foods (Bourne, 1982). We used an Instron (model 4301) materials testing machine (Instron, Canton, Mass.) to apply penetrometer tests to fruit. The speed of penetration was set at 240 mm·min⁻¹, and the test was stopped after penetration to 8 mm deep.

The twist tester consists of a blade fixed at 90° to a horizontal spindle (Studman and Yuwana, 1992) (Fig. 1d). The fruit is pushed onto the spindle and rotated around the spindle until the blade crushes the flesh and the fruit turns freely. The resisting moment is generated by a counterweight attached to an offset arm. The angle that the fruit rotated before tissue failure was measured using a potentiometer connected to a chart recorder (model R-01; Rikadenki, Tokyo). This angle then was converted into a crushing strength using formulas described by Studman and Yuwana (1992).

In 1991, we compared Effegi hand-held penetrometer, Effegi penetrometer mounted in a motorized press, EPT pressure tester, and Instron. In 1992, we compared Effegi hand-held penetrometer, Effegi penetrometer mounted in a drill press, Instron, and twist test. All measurements were made to the nearest 1 N, except for the twist test, which was measured to the nearest 7 KPa.

Operators. All operators were members of the postharvest group of HortResearch and had at least 1 year of prior experience using hand-held Effegi penetrometers. None of the operators had used the EPT pressure tester, twist tester, or a drill press to operate the Effegi penetrometer before these experiments. Operator were allowed sufficient measurements to familiarize themselves with the devices before the start of each series of measurements. Pseudonyms identify operators.

Experimental design. Slices of skin were removed at four equidistant points around the equator of each fruit. Firmness then was measured on opposite sides of the fruit for each device–operator combination before passing the fruit to the next device–operator combination for firmness measurements on the remaining opposite sides. This process allowed us to compare the device–operator combinations relative to variation between different measurements on the same fruit, thus, excluding any fruit-to-fruit component of variability. Each person (Amy, Cloe, Dan, and Ed in 1991 and Ben, Dan, Ed, and Fay in 1992) used each instrument with the exception of the Instron and motorized press, which were presumed to be operator independent and, thus, were operated by F.R.H.

Two fruit were used, giving two sets of measurements for each pair of device–operator combinations (Table 1). Formally, the design was a balanced incomplete block design, with two devices per block (fruit) (Gacula and Singh, 1984; Pearce, 1983). There were two results for each device in each block. A single fruit constituted a "block."

Statistical analysis. From examining plots of residuals against fitted values in the analysis of variance (ANOVA) calculations described later in this paper, we concluded that using log (firmness) would best ensure that variances were comparable over the whole range of firmness. To reflect the experimental design correctly, we needed an ANOVA table (Table 2) in which there were three major strata, or levels, of variability. These were between fruit ("fruit" stratum), between different pairs of measurements on the same fruit ("fruit × device" stratum), and within pairs of measurements made by one device–operator combination ("fruit × device × units" stratum). Relatively imprecise information on the comparison between device–operator combinations appears in the interfruit (or interblock) analysis of the fruit stratum. More precise information on this comparison appears in the intrafruit (or intrablock) analysis of the fruit × device stratum. For a discussion of the interblock and intrablock analysis for a balanced incomplete design, see for example Gacula and Singh (1984). The final or fruit × device × units stratum contains no information on the comparison between device–operator combinations. The sum of squares in this stratum (180 × 0.0085) is the total of the sums of squares about the mean for the two results for each device–operator combination used in each of the 90 fruit.

For our analysis, we used the Genstat ANOVA routine

Table 1. An example showing the type of experimental design used to compare penetrometers. The table indicates how four operator–device combinations were tested against each other. For each operator–device combination, firmness was measured twice, one on each of two opposing sides of the fruit. In this example, there are six pairwise comparisons, against 45 and 65 pairwise comparisons in actual experiments in 1991 and 1992, respectively.

Operator–device combination	Operator–device combination			
	Dan[z]-Hand[y]	Dan-EPT	Cloe-Hand	Cloe-EPT
Dan-hand		1[x], 2	3, 4	5, 6
Dan-EPT			7, 8	9, 10
Cloe-hand				11, 12
Cloe-EPT				

[z]Operators are identified by pseudonyms.
[y]Instruments are identified by abbreviations: hand = hand-held Effegi penetrometer and EPT = EPT pressure tester.
[x]Each number identifies an individual fruit.

Table 2. Overall analysis of variance table for stored kiwifruit in 1991. The Genstat statement BLOCK fruit/device gives the error structure needed to get this table in which each analysis of variance mean square is associated with its appropriate residual or "error" mean square. The residual mean square is a variance estimate for variation in that stratum.

Source of variation	Degrees of freedom	Mean square
Fruit stratum (interfruit analysis)		
Device	9	2.233
Residual	80	0.127
Fruit × device stratum (intrafruit analysis)		
Device–operator	9	3.146
Residual	81	0.0096
Fruit × device × units stratum	180	0.0085
Total	359	

J. Amer. Soc. Hort. Sci. 121(5):927–936. 1996.

929

(Maindonald, 1992; Payne et al., 1993; Pearce, 1983), which is able to give the table described in the previous paragraph. Table 2 may be obtained from an ANOVA table that ignores the hierarchical error structure by a suitable repartitioning and grouping of the entries in the rows. Typically, it is expected that the error mean square for differences between fruit (i.e., in what Table 2 calls the fruit stratum) will be larger than the error mean square for comparisons between different devices on the same fruit (i.e., in the fruit × device stratum. By using a design and accompanying analysis that allows comparison of devices relative to this second and smaller error mean square, we avoid the unnecessary "noise" that differences between fruit would otherwise add to the comparison of devices. This analysis may be extended to split the stratum variances into components of variance, as in the Genstat REML analysis (Payne et al., 1993). We have not taken this further step as it would not assist data interpretation.

ANOVA tables, tables of means, and standard errors of differences (SEDs) of means were determined for each combination of year, before or after storage, and fruit (kiwifruit or apple). It was necessary to make comparisons between storage and harvest results relative to between-fruit variation.

To get an approximate overall comparison between device–operator combinations, we made the simplifying assumption that within-fruit variances for the different device–operator combinations were similar. The overall SEDs and least significant differences that are presented should be used as broad indicators of differences. The variance ratios for testing for differences between device–operator combinations always were so large that formal use of an F test was unnecessary.

In addition, analyses were performed for each individual device–operator combination. As before, these partitioned the total sum of squares into a sum of squares for between-fruit variation and two sets of sums of squares for within-fruit variation. The

mean square (variance estimate = sum of squares ÷ degrees of freedom) based on differences between measurements on opposite sides was used as a measure of inherent variability (lack of precision) of each device–operator combination. Boxplots were used to display fruit-to-fruit variation and differences of individual measurements for a device–operator combination from the fruit mean. These helped to identify outliers. We used the robloc function in S-PLUS (Statistical Sciences, 1991) to determine robust estimates for the variance based on differences between measurements on opposite sides. Such estimates are not affected by an occasional large outlier.

We used Bartlett's test for homogeneity of variance (Miller, 1986) to check on apparent variance differences between device–operator combinations. Because Bartlett's test is sensitive to nonnormality, it was important to use this test with checks on outliers and with a low significance level ($P = 0.01$).

Analysis of twist tester measurements. The data presented in Fig. 2 allowed us to relate kiwifruit twist test measurements (1992 only) to penetrometer measurements. To compare variability between the two types of device, we needed to reduce them to a common scale. Because we wished to place the two scales on an equal footing, neither of the regressions log y on log x nor log x on log y was appropriate for this purpose. Instead, we estimated a functional relationship, as described in Kendall and Stuart (1979, p. 409–412). We assumed that with y = twist test result and x = penetrometer result, the errors in log y and log x have equal variance. Then log y = 2.428 + 1.164 (SE = 0.025) log x. The line is shown in Fig. 3. The same transformation has been used for apple twist measurements to get comparability with other devices.

Over the range of measurements used, the relationship was linear. The approach can be extended to handle situations where the relationship is nonlinear. For example, one might precede the calibration just described with using a smoothing function to
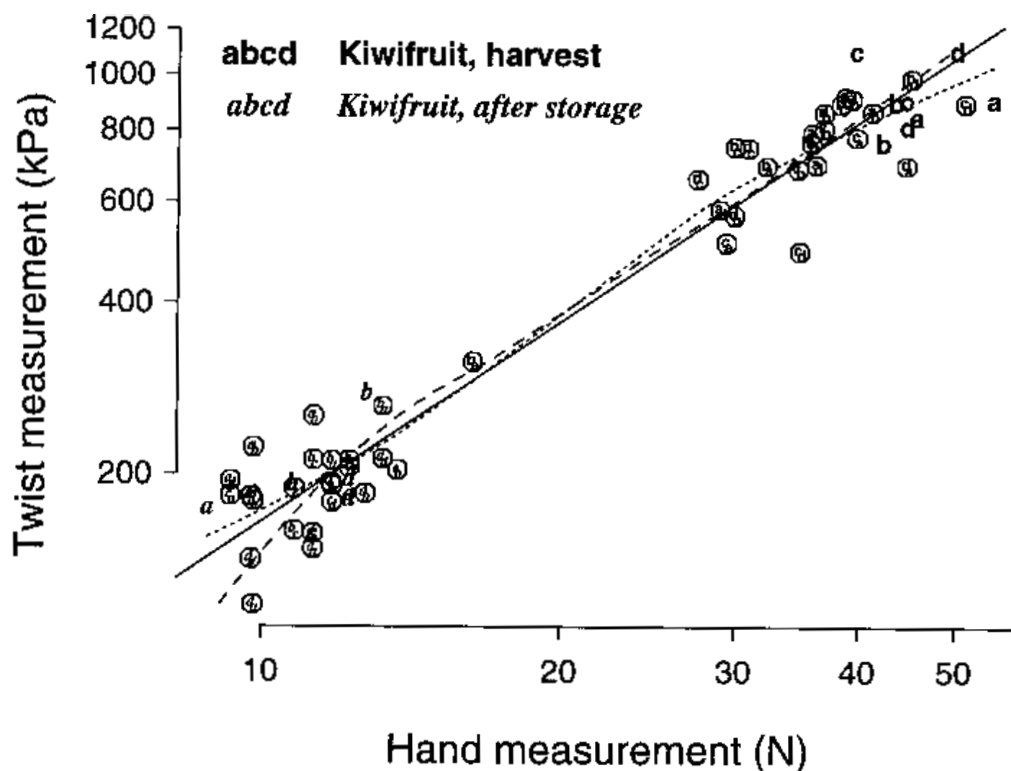

Fig. 2. Relationship between penetrometer and twist test measurements of kiwifruit firmness. Measurements obtained using the twist test are plotted against firmness according to penetrometer measurements. The solid line is obtained as described in the text. Broken lines (······ and ----) are smoothed curves using S-PLUS loess smoothing function for y on x and for x on y, respectively.
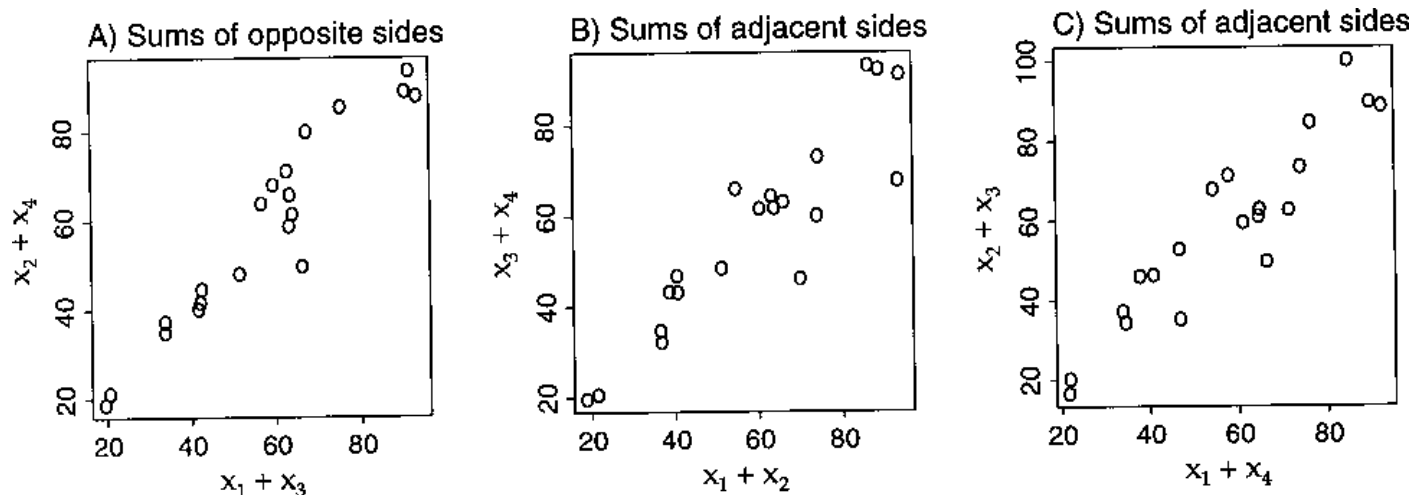
Fig. 3. Relationship between firmness of opposite and adjacent quarters of kiwifruit. Firmness was measured using an Instron with a probe 7.9 mm in diameter. Measurements x_1, x_2, x_3, and x_4 were made in order around the fruit equator. The graphs show (**A**) sums of pairs of opposite sides plotted against each other and (**B** and **C**) sums of pairs of adjacent sides plotted against each other. Values are in newtons.

provide a transformation to a linear relationship. The S-PLUS "loess" smoothing function (Chambers and Hastie, 1991) is suitable for this purpose. The presentation of data in the figures is based on user styles that were discussed and recommended by Maindonald (1992).

Results and Discussion

Preliminary investigations. There were four measurements on each fruit, two (on opposite sides) with one device–operator combination and two with another device–operator combination. This methodology was determined after preliminary studies with apples and kiwifruit using one device (Instron). There are three comparisons of penetrometer measurements x_1, x_2, x_3, x_4 (taken in order around the fruit equator). One ($x_1 + x_3 - x_2 - x_4$) compares pairs of opposite measurements, while the other two ($x_1 + x_2 - x_3 - x_4$) and ($x_2 + x_3 - x_1 - x_4$) compare pairs of adjacent measurements. The variance of these comparison statistics measures the consistency of the pairs whose difference is taken, with small variance indicating high consistency. The variances were 28.3, 42.2, and 37.5 for apples and 42.2, 85.6, and 70.5 for kiwifruit. While the differences were not significant ($P = 0.05$), and outliers were a contributing influence, these results provided a prima facie case for working with pairs of opposite measurements. The plots shown in Fig. 3 provide a useful visual check on the consistency between different pairs of readings. The first plot, comparing pairs of opposite measurements, shows less scatter than the remaining plots, which compare pairs of adjacent measurements.

Fruit variability. Within-fruit mean squares (variances) for different devices on the same fruit in Table 3 gives an indication of

Table 3. Comparison of between- and within-fruit (fruit × device stratum) mean squares. Notice that between-fruit mean squares generally are much larger than within-fruit mean squares. This difference has implications for experimental design. The mean squares have been used to calculate the estimates of the number of fruit required to achieve the same precision when different device–operator combinations are used with different fruit that appear in the last column.

Fruit		Mean squares			Fruit no.	
		Between	Within	Eff.[z]	No.	Alt. expt.[y]
		1991				
Kiwifruit	Harvest	0.40 (80)[x]	0.024 (81)[x]	0.556	90	90 × 9.3
	Storage	0.13	0.0096	0.556	90	90 × 7.3
Apple	Harvest	0.0068	0.0026	0.556	90	90 × 1.5
	Storage	0.014	0.0021	0.556	90	90 × 3.7
		1992				
Kiwifruit	Harvest	0.15 (143)[x]	0.015 (144)[x]	0.542	156	156 × 5.6
	Storage	0.057	0.018	0.542	156	156 × 1.7
Apple	Harvest	0.0084	0.0042	0.542	156	156 × 1.1
	Storage	0.015	0.0031	0.542	156	156 × 2.6

[z]Eff. = efficiency of estimation of means of device–operator combinations in the within-fruit stratum.
[y]Alt. expt. = number of fruit required to achieve the same precision in an alternative experimental design that tests different device–operator combinations at the between-fruit stratum with four measurements per fruit. Thus, for 1991, kiwifruit at harvest, the relative number is (0.556 × 0.40) ÷ 0.024 = 9.3.
[x]Degrees of freedom.

differences in variability between kiwifruit and apples and between harvest and storage. These are average variabilities over devices and over operators. Table 3 also includes information that allows an assessment of the benefit of our design relative to a design in which different device–operator combinations were always tested on different fruit. To keep the comparison fair, we assume, for both designs, four measurements per fruit. The table shows how, depending on fruit-to-fruit consistency, an experimental design that allows comparisons to be made at the within-fruit level of variation may reduce the number of fruit and measurements needed to achieve a given precision dramatically. Using a better design often is preferable to increasing the number of experimental units. Where fruit were more consistent (compare apples with kiwifruit; Table 3) the gains are smaller.

Differences between methods. Measurements of firmness are used in two ways. First, the absolute firmness value can be used in a regulatory fashion to define the minimum standard of fruit that

J. Amer. Soc. Hort. Sci. 121(5):927–936. 1996.

931

Fig. 4. Fruit firmness according to measurements by Instron, EPT pressure tester, and Effegi penetrometer operated by hand or using a motorized press. Differences between operators are indicated for the hand-held Effegi penetrometer and EPT pressure tester. Measurements were made during the 1991 fruit season using apples (**A**) at harvest and (**B**) after storage and kiwifruit (**C**) at harvest and (**D**) after storage. EPT = EPT pressure tester, Hand = hand-held Effegi penetrometer, and M. Press = motorized press.

A) Apples at Harvest

Fay Drill — Dan Hand — Ben Drill — Fay Hand — Ben Hand — Ed Drill — Dan Drill — Instron — Ed Hand

72   73   74   75   76   77   78   79

├──────────────────────────────────┤ LSD (Tukey)

B) Apples After Storage

Fay Drill — Dan Drill — Instron — Ed Drill — Ben Drill — Fay Hand — Dan Hand — Ed Hand — Ben Hand

65   66   68   70   72   73

├──────────────────────────────────┤ LSD (Tukey)

C) Kiwifruit at Harvest

Instron — Dan Twist — Dan Hand — Ed Twist — Fay Hand — Dan Drill — Ben Twist — Fay Drill — Fay Twist — Ed Hand — Ed Drill — Ben Drill — Ben Hand

28   30   32   34   36   38   40

├──────────────┤ LSD (Tukey)

D) Kiwifruit After Storage

Instron — Ben Twist — Ben Hand — Ed Hand — Dan Twist — Fay Hand — Fay Drill — Dan Hand — Ed Twist — Fay Twist — Ed Drill — Dan Drill — Ben Drill

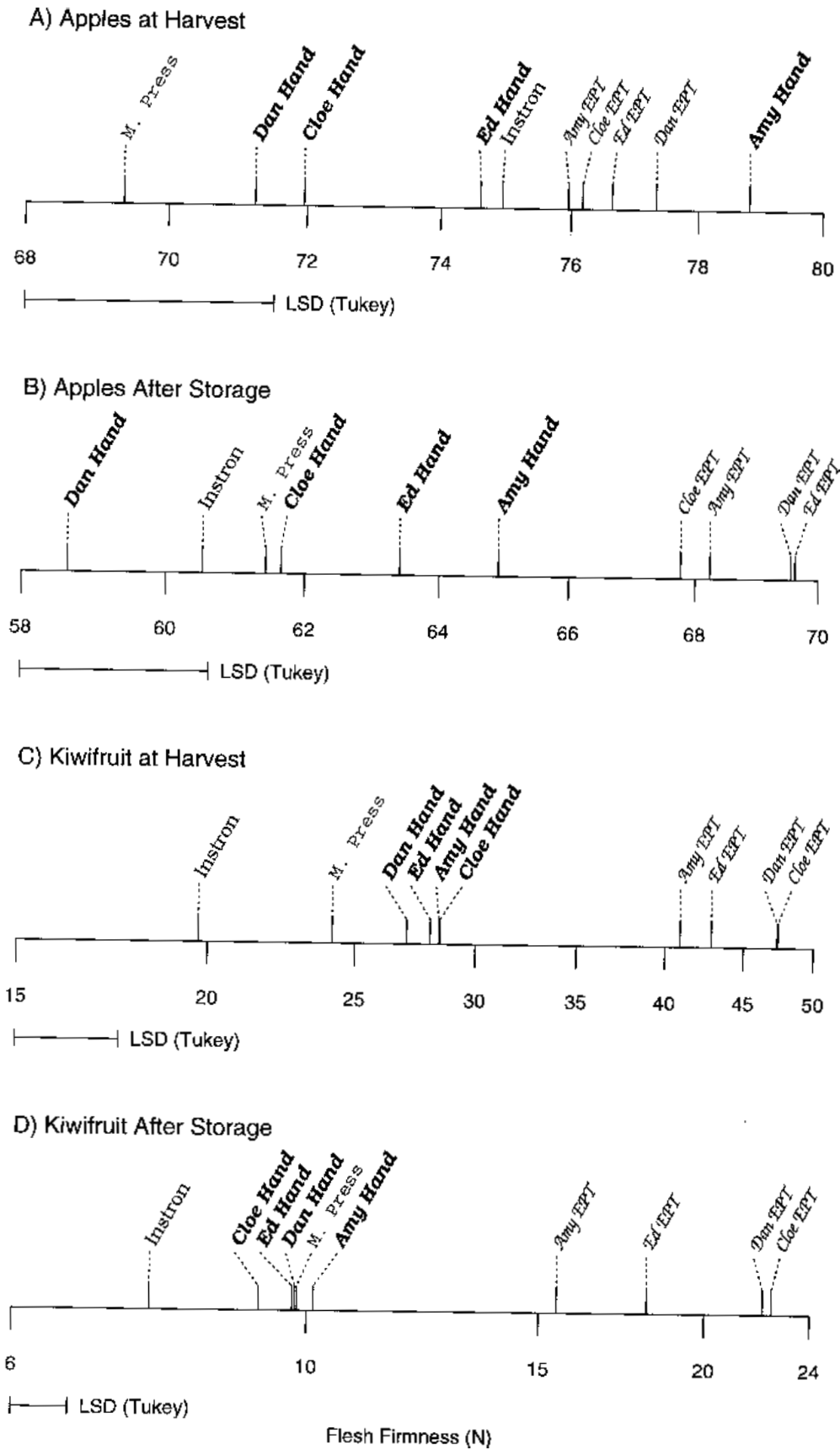8   9   10   11   12   13   14

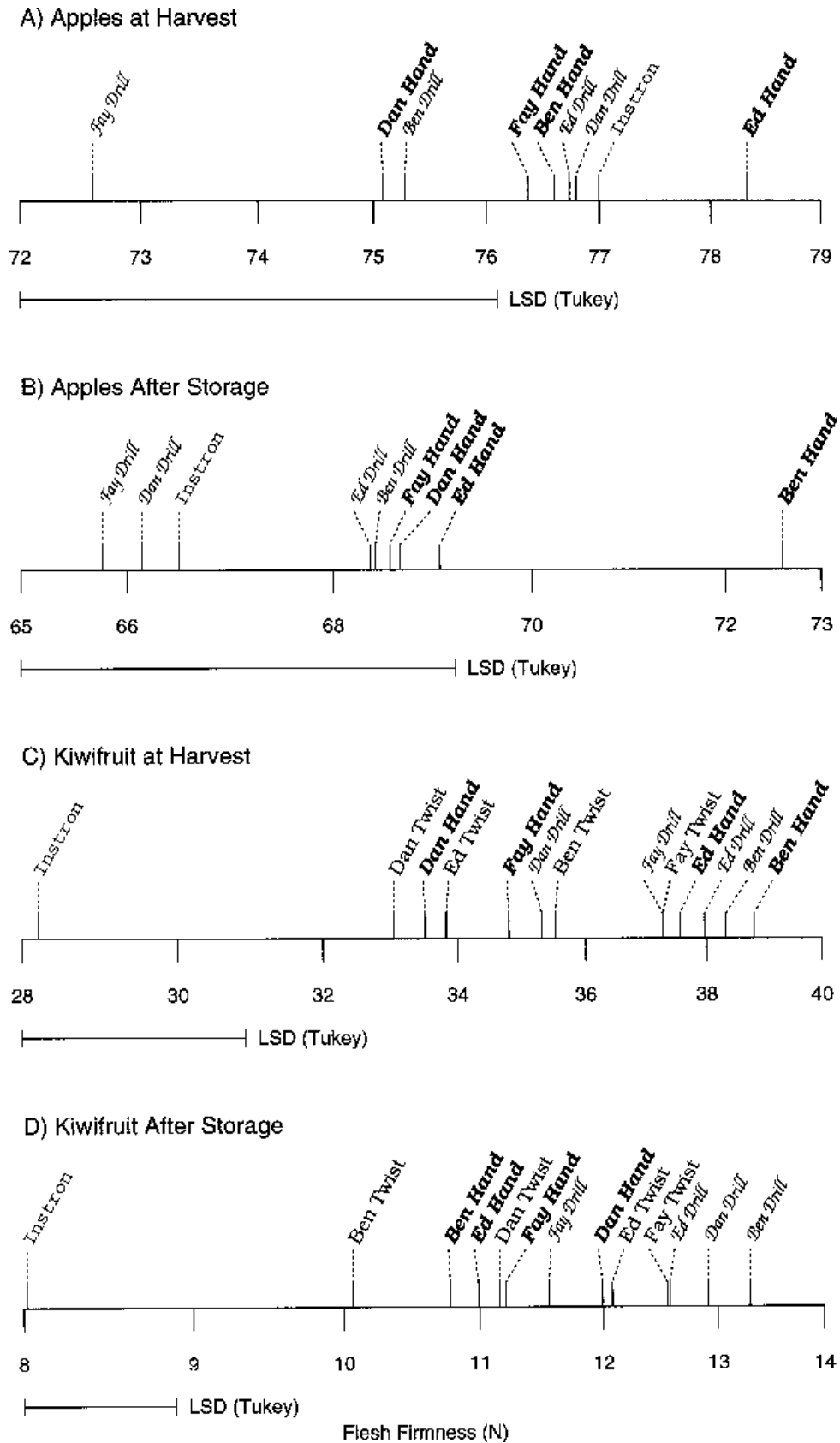├──────────┤ LSD (Tukey)

Flesh Firmness (N)

Fig. 5. Fruit firmness according to measurements by Instron, twist test, and Effegi penetrometer operated by hand or using a drill press. Differences between operators are indicated for the twist test (except with apples, see text) and Effegi penetrometer operated by hand or using a drill press. Measurements were made during the 1992 season using apples (**A**) at harvest and (**B**) after storage and kiwifruit (**C**) at harvest and (**D**) after storage. Hand = hand-held Effegi penetrometer, Drill = Effegi penetrometer operated using a drill press, and Twist = twist test.

growers must achieve before their fruit will be accepted by export or retail organizations. Second, change in firmness is used to characterize the rate at which fruit soften. In our study, measurements of firmness varied according to the device used. With hard fruit (apples at harvest, Fig. 4a and 5a) all devices except the motorized press provided similar estimates of firmness once differences between operators were excluded. The motorized press (Fig. 4a) underestimated firmness. With soft fruit (stored apples, 60 to 69 N and kiwifruit at harvest and after storage) (Fig. 4), the EPT pressure tester provided the highest estimate of firmness followed by the hand-held Effegi penetrometer, motorized press, and the Instron, respectively. Bongers (1992), who assessed firmness of apples (average firmness between 50 and 58 N, which is equivalent to soft apples in our study), using the EPT pressure tester, Instron, and Effegi penetrometer mounted in a drill press found a similar ranking of devices. The results from both studies indicate the need for standardization in any regulatory use of firmness measurements. Differences in machinery and operator may considerably affect the measurement of absolute firmness.

Ability to discriminate between hard and soft fruit. Examining Fig. 4 and 5 may suggest that the Instron is better than other devices at detecting differences in firmness. For example, the Instron gave a 14.8-N decline in apple firmness in 1991 (Fig. 4a; cf. Fig. 4b). Measurements by EPT pressure tester and Effegi penetrometer operated by hand or motorized press indicated average firmness declined an average of 9.3, 11.2, and 8.1 N, respectively. However, while the firmness difference is larger for the Instron, this difference was often offset by the lower precision with which this difference is measured. Differences between measurements on opposite sides of one fruit provide one possible measure of the inherent precision of equipment. We checked variance estimates for the possible effect of outliers. Table 4 gives nonrobust and, where these differed by >15%, robust variance estimates. In general, the device for which the variance is smallest is clearly best. Outliers, which may be due to occasional soft or hard spots, should ordinarily be excluded in the comparison. The robust estimates effectively exclude outliers. Frequent outliers, however, may indicate that the device is prone to giving aberrant readings.

Three of eight Instron variance estimates were ≈0.03, and much larger than variances 0.0017 to 0.018 estimated for other devices (Table 4). Using robust estimates reduced all these estimates by a factor of three or more. The Instron may have been prone to occasional aberrant readings due to the way fruit were held by hand (perhaps the fruit twisted during penetration of the probe). Using S-PLUS's (Statistical Sciences, 1991) default boxplot criterion for outliers, average numbers of outliers per fruit in differences between the two reading were Instron 9%, motorized press 8%, hand-held Effegi penetrometer 3%, Effegi penetrometer operated by drill press 6%, twist test 5%, and EPT pressure tester 6%. A χ^2 test suggests that the outlier rate does differ between devices ($P = 0.05$). For each of the columns in Table 4, the Instron robust variance estimate was always largest or close to largest.

Comparing fruit after harvest and after storage. A comprehensive comparison would take into account the effect of operator-to-operator variability and assess ability to discriminate when presented with a wide range of firmness. One such test is the ability to discriminate between fruit at harvest and after storage. For this comparison, we divided differences in firmness estimates between harvest and storage by a SED that was based on fruit-to-fruit variability (i.e., on the between-fruit mean square of Table 3). This gives a t statistic, which we used as a measure of how well the device discriminates. For comparing t statistics, a difference of ≈2 is significant at $P = 0.05$, and a difference of 2.8 is significant at $P = 0.01$. Results from this analysis are presented in Table 5.

The most precise measurements of change in firmness according to calculated t statistics were obtained using the Instron and hand-held penetrometer for apples in 1991, the Instron for apples in 1992, the hand-held penetrometer for kiwifruit in 1991, and the twist test for kiwifruit in 1992. Note that the hand-held penetrometer was among the more precise devices. The benefit of the Instron may be that it is independent of operator.

Influence of operator differences. The importance of operator differences should be considered relative to the absolute changes in firmness that are likely to occur during storage, the relative firmness differences that are likely to occur between treatments, or both. Thus, differences between operators of ≤2.16 N associated with hand-held Effegi penetrometers (Fig. 4d) are of relatively minor importance when compared to the 85-N decline in firmness that can occur during kiwifruit softening (Beever and Hopkirk, 1990). However, apples soften only slightly during storage and differences between operators of

Table 4. Within-fruit variance estimates. Devices that give small variances, and hence are more precise, are to be preferred. Where the robust estimate of variance is >15% less than the nonrobust estimate, this is given also. Robust estimates omit any "outliers" from the calculation.

Instrument	Harvest		Storage	
	1991	1992	1991	1992
Kiwifruit				
Instron	0.013 (18 df)	0.030 (24 df) (0.010)[z]	0.034 (18 df)	0.0095 (24 df) (0.0095)[z]
Motorized press	0.018 (18 df) (0.011)[z]	---	0.0053	---
Hand-held	0.0067 (72 df)	0.010 (96 df)	0.0051 (81 df) (0.0049)[z]	0.0067 (105 df) (0.0038)[z]
EPT pressure tester	0.0072 (72 df)	---	0.0064 (0.0049)[z]	---
Drill press	---	0.011 (96 df)	---	0.018 (0.011)[z]
Twist test	---	0.0054 (96 df)	---	0.0084
Apple				
Instron	0.0033	0.0035	0.0035	0.030 (0.0073)[z]
Motorized press	0.0022	---	0.0024 (0.00090)[z]	---
Hand-held	0.0024	0.0021	0.0017 (0.0013)[z]	0.0024
EPT pressure tester	0.0025	---	0.0018	---
Drill press	---	0.0034	---	0.0040
Twist test	---	0.0025	---	0.0028

[z]Robust estimate of variance.

≤7.6 N are relatively large when compared to the 14.4-N decline in firmness (from 74.9 to 60.5 N; Instron, Fig. 4 a and b).

To eliminate operator differences, it is recommended that the same operator be used for all penetrometer measurements (Blanpied et al., 1978). The data presented in our study (Fig. 4 and 5) and in earlier studies (Blanpied et al., 1978; Voisey, 1977) confirm that using different operators can significantly influence firmness values. Bongers (1992), however, did not detect any difference between the four operators used in his study.

All hand-operated devices examined in this study were subject to operator differences, with the magnitude of the difference varying between devices and between soft and hard fruit (Figs. 4 and 5). Operator differences associated with the EPT pressure tester were minimal on firm apples but more substantial on soft apples and kiwifruit (Fig. 4). With hard apples, operator differences were greater when Effegi penetrometers were operated using a drill press than when operated by hand (Fig. 5a). This ranking reversed with soft apples and kiwifruit (Fig. 5 b–d). We found that the drill press was more awkward and slower to use, and did not always reduce operator differences. Before proceeding with using a drill press, it is worth checking that this method is providing some advantage over the hand-held penetrometer. Different results might be expected depending on the model of drill press used and on the strength of the springs and smoothness of the downward action.

The cause of operator differences associated with hand-held penetrometers may arise from the speed, angle, depth, or acceleration used to push the probe into the fruit. In preliminary investigations using the Instron, we found that the speed and angle at which the probe enters the fruit appears to have little effect (data not shown). This is in agreement with recommendations that the Instron can operate at speeds between 50 and 250 mm·min^{-1} when measuring flesh firmness (Smith, 1985). Examining force–distance curves generated by the Instron (data not shown) indicated that force reached a maximum value within the first few millimeters of penetration in this study and, subsequently, declined or remained constant (type B and C curves; Bourne, 1982). Thus, operator differences are unlikely to be caused by differences in the depth that the probe is pushed into the fruit.

Comparisons of measurements by the motorized press and Instron suggest that operator differences may be associated with the way the probe was accelerated into the fruit. The motorized press was subject to a systematic error. While the penetrometer was driven downward at a constant speed, acceleration into the fruit was determined by the distance the spring was compressed before the probe entered the fruit. Thus, we speculate that acceleration into the flesh was slow in hard fruit and rapid in soft fruit. Comparisons of the motorized press with the Instron, which has a constant velocity, indicate that the motorized press underestimated firmness of hard fruit (Fig. 4a) and overestimated firmness of soft fruit (Fig. 4d). Voisey (1977) found that, while most operators apply a smooth increase in force when using hand-held penetrometers, other operators changed the rate of force application or momentarily relaxed the force just as the tissue yielded.

Ranking of operators according to firmness measurements gave a similar order on each assessment date. For example, Amy always provided the highest firmness measurement with a hand-held Effegi penetrometer (Fig. 4). Thus, it may be reasonable to assume that operator-to-operator differences are similar at all evaluation times. Further studies are required to validate this finding. Our experiments involved relatively small numbers of measurements for any operator. We have no information whether operators can maintain consistency over a long series of measurements. Operator

Table 5. Precision with which devices discriminate between hard and soft fruit. Values represent t statistics, which in this context, we use to measure how well the device discriminates between fruit at harvest and fruit after storage. The higher the t statistic the better the precision. See text for full description.

| | Harvest | |
Instrument	1991	1992
Kiwifruit		
Instron	6.1[z]	19.2
Motorized press	10.7	---
Hand-held	9.6–16.6[y]	18.9–28.4
EPT pressure tester	10.9–15.0	---
Drill press	---	19.1–26.1
Twist test	---	25.1–26.1
Apple		
Instron	8.7	13.5
Motorized press	6.2	---
Hand-held	5.2–11.2	5.4–9.6
EPT pressure tester	5.6–7.8	---
Drill press	---	3.6–6.5
Twist test	---	–4.1–5.1

[z]Single t-statistics are given where device is independent of operator.
[y]The range of t statistics is given when multiple operators have used the device.

consistency is an interesting issue, particularly if measurements of firmness by a single operator vary from day to day depending on his or her mood and health.

Ease of use of devices. Hand-operated devices such as the Effegi penetrometer give full manipulative control of position of penetrometer and fruit. When the Effegi was mounted in a drill press and a motorized press, the operation was slowed by the requirement for accurate placement of the fruit. All Effegi penetrometer measurements involved reading the dial and entering the value onto paper or directly into a personal computer. This increases the possibility of errors occurring during reading and recording of values. In our opinion, the EPT pressure tester was the easiest device to use to collect firmness data. Fruit were easily positioned under the probe, the movement of the hinged arm was simple, and the data was automatically transferred to a computer or printed onto paper.

When considering the ease of use, it is important to examine how the measurement of firmness fits in with other tests of fruit quality. For example, evaluating apple maturity and quality at harvest routinely involves measurements of internal ethylene concentration, skin background color, flesh firmness, soluble solids concentration, and starch pattern index (e.g., Watkins et al., 1992). In our laboratory, we often combine measurements of firmness and soluble solids by transferring juice expressed onto the tip of the penetrometer probe directly onto a refractometer (Watkins et al., 1992). This operation is easily performed when using the Effegi penetrometer operated by hand, Effegi penetrometer mounted in a drill press, and the EPT pressure tester.

Conclusion

This study shows the large gains that are available from an effective experimental design. It demonstrates that differences between operators can be a significant problem when hand-operated devices are used to measure firmness. Our results reinforce the recommendation that the same operator be used when hand-operated penetrometers are used to measure flesh firmness

J. Amer. Soc. Hort. Sci. 121(5):927–936. 1996.

935

for comparisons of pre- and postharvest treatments or to follow fruit softening (Blanpied et al., 1978). When devices, such as drill presses, are used to drive the penetrometer probe into fruit, it would be sensible to check that they do indeed improve the precision of measurement and reduce differences between operators over the more simple devices available. Using a materials testing machine to drive the penetrometer probe into fruit seemed to provide a consistent method of measuring firmness. With apples the Instron provided a precise measurement of softening. Although the Instron provided less precise measurements of kiwifruit softening, it has the benefit of greater operator independence.

Literature Cited

Abbott, J.A., A.E. Watada, and D.R. Massie. 1976. Effegi, Magness-Taylor, and Instron fruit pressure testing devices for apples, peaches, and nectarines. J. Amer. Soc. Hort. Sci. 101:698–700.

Beever, D.J. and G. Hopkirk. 1990. Fruit development and fruit physiology, p. 97–126. In: I.J. Warrington and G.C. Weston (eds.). Kiwifruit: Science and management. N.Z. Soc. Hort. Sci., Auckland.

Blanpied, G.D., W.J. Bramlage, D.H. Dewey, R.L. LaBelle, L.M. Massey, Jr., G.E. Mattus, W.C. Stiles, and A.E. Watada. 1978. A standardized method for collecting apple pressure test data. New York's Food and Life Sciences Bul. 74:1–8.

Bongers, A.J. 1992. Comparison of three penetrometers used to evaluate apple firmness. Washington State Tree Fruit Postharvest J. 3:7–9.

Bourne, M.C. 1982. Food texture and viscosity: Concept and measurement. Academic, San Diego.

Chambers, M.C. and T.J. Hastie. 1991. Statistical models in S. Wadsworth and Brooks/Cole, Pacific Grove.

Gacula, M.C. and J. Singh. 1984. Statistical methods in food and consumer research. Academic, Orlando, Fla.

Kendall, M.G. and A. Stuart. 1979. The advanced theory of statistics, 4th ed. Griffin, London.

Magness, J.R. and G.F. Taylor. 1925. An improved type of pressure tester for the determination of fruit maturity. U.S. Dept. Agr. Dept. Circ. no. 350.

Maindonald, J.H. 1992. Statistical design, analysis, and presentation issues. N.Z. J. Agri. Res. 35:121–141.

Miller, R.G., Jr. 1986. Beyond ANOVA, basics of applied statistics. Wiley, New York.

Payne, R.W., P.W. Lane, P.G.N. Digby, S.A. Harding, P.K. Leech, G.W. Morgan, A.D. Todd, R. Thompson, G. Tunnicliffe Wilson, S.J. Welham, and R.P. White. 1993. Genstat 5 Release 3 reference manual. Clarendon Press, Oxford.

Pearce, S.C. 1983. The agricultural field experiment: A statistical examination of theory and practice. Wiley, New York.

Saltveit, M.E., Jr. 1978. Selecting an experimental design for pressure testing apples. J. Amer. Soc. Hort. Sci. 103:105–109.

Smith, S.M. 1985. Measurement of the quality of apples: Recommendations of an EEC working group. Commission European Comm.

Statistical Sciences. 1991. S-PLUS reference manual, version 3.0. Statistical Sciences, Seattle.

Studman, C.J. and Yuwana. 1992. Twist test for measuring fruit firmness. J. Texture Studies 23:215–227.

Voisey, P.W. 1977. Examination of operational aspects of fruit pressure tests. Can. Inst. Food Sci. Technol. 10:284–294

Watkins, C. and J. Harman. 1981. Use of penetrometer to measure flesh firmness of fruit. Orchardist N.Z. 54:14–16.

Watkins, C.B., I.J. Horner, and I.B. Ferguson. 1992. The effects of collar rot (*Phytophthora cactorum*) on "Cox's Orange Pippin" fruit quality. J. Hort. Sci. 67:495–500.

Worthington, J.T. and J.N. Yeatman. 1968. A statistical evaluation of objective measurement of apple firmness. Proc. Amer. Soc. Hort. Sci. 92:739–747.

936

J. Amer. Soc. Hort. Sci. 121(5):927–936. 1996.